9-1-2016

Ancient DNA Investigation of a Medieval German Cemetery Confirms Long-Term Stability of CCR5-Δ32 Allele Frequencies in Central Europe

Abigail Bouwman
University of Zurich, abigail.bouwman@iem.uzh.ch

Natallia Shved
University of Zurich, nshved@yahoo.com

Gulfirde Akgul
University of Zurich, onder_guli@hotmail.com

Frank Ruhli
University of Zurich, frank.ruehli@iem.uzh.ch

Christina Warinner
University of Oklahoma, warinner@shh.mpg.de

Title

Ancient DNA investigation of a medieval German cemetery confirms long-term stability of CCR5-Δ32 allele frequencies in Central Europe

Authors

Abigail Bouwman[1], Natallia Shved[1], Gülfirde Akgül[1], Frank Rühli[1], Christina Warinner[1,2]

Affiliations

[1]Institute of Evolutionary Medicine, University of Zürich, Zürich, CH-8057, Switzerland

[2]Department of Anthropology, University of Oklahoma, Norman, OK 73019, USA

Keywords

Ancient DNA; CCR5; medieval

Running Title

CCR5 allele frequency in medieval Germany

Corresponding Author:

Christina Warinner

Department of Anthropology

University of Oklahoma

101 David L. Boren Blvd.

Norman, OK 73019, USA

christina.warinner@ou.edu

Date Submitted: October 7 2016

Date Accepted: February 12 2017

Abstract

The *CCR5-Δ32* mutation present in European populations is among the most prominently debated cases of recent positive selection in humans. This allele, a 32-bp deletion that renders the T-cell CCR5 receptor non-functional, has important epidemiological and public health significance, as homozygous carriers are resistant to several HIV strains. However, although the function of this allele in preventing HIV infection is now well described, its human evolutionary origin is poorly understood. Initial attempts to determine the emergence of the *CCR5-Δ32* allele pointed to selection during the 14th century Black Death pandemic; however, subsequent analyses suggest that the allele rose in frequency more than five thousand years ago, possibly through drift. Recently, three studies have identified populations pre-dating the 14th century CE that are positive for the *CCR5*-Δ*32* allele, supporting the claim for a more ancient origin. However, these studies also suggest poorly understood regional differences in the recent evolutionary history of the *CCR5*-Δ*32* allele. Here a new hydrolysis probe-based Real-Time PCR (RT-PCR) assay was designed to ascertain *CCR5* allele frequency in 53 individuals from a 10th-12th century CE church and convent complex in central Germany that predates outbreaks of the Black Death pandemic. High confidence genotypes were obtained for 32 individuals, and the results of the study show that *CCR5*-Δ*32* allele frequency has remained unchanged in this region of Central Europe over the last millennium, suggesting that there has been no strong positive selective pressure over this time period and confirming a more ancient origin for the allele.

Acknowledgments: This research was supported by the Mäxi Foundation Zürich and the Novartis Foundation. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

The chemokine receptor CCR5 is expressed on the surface of host immune cells and, together with the CD4 receptor, serves as the binding and entry point of HIV. Non-functionality of the CCR5 receptor on CD4-bearing T lymphocytes disrupts the formation of the CD4-CCR5-HIVglycoprotein-120 binding complex that is required for infection of non-synctium-inducing (NSI) HIV strains (Michael et al. 1997). *CCR5*-Δ*32*, a genetic variant corresponding to a 32-bp deletion in the CCR5 gene, results in a non-functioning chemokine receptor. *CCR5*-Δ*32* homozygosity imparts resistance to NSI HIV strains, and heterozygosity results in delayed HIV acquisition and slower HIV progression to AIDS (Dean et al. 1996; Huang et al. 1996; Liu et al. 1996; Samson et al. 1996). Today, *CCR5*-Δ*32* allele frequency is approximately 1-15% in Europe, Russia, the Middle East, and the Indian subcontinent, but rare or absent elsewhere (Martinson et al. 1997). The high frequency, limited biogeographic distribution, and difficulty in dating the *CCR5*-Δ*32* allele has led to intense debate over whether or not current genotype frequencies arose by selection or drift (Hedrick and Verrelli 2006).

Although HIV infection of humans does not predate the 20th century and thus cannot explain current *CCR5*-Δ*32* allele patterns (Sharp and Hahn 2011), it has been hypothesized that other infectious agents, such as *Yersinia pestis* (plague) bacteria (Stephens et al. 1998), the smallpox variola virus (Galvani and Slatkin 2003), or other unknown human pathogens, may involve the CCR5 receptor in infection, thereby offering a potential mechanism for positive selection in the past. However, the involvement of the CCR5 receptor in these infections is ambiguous and contested. For example, *CCR5*-Δ*32* does not appear to be protective against *Yersinia* spp. infection or mortality in mouse models (Mecsas et al. 2004), but it does affect *Y. pestis* uptake rates by macrophages (Elvin et al. 2004). With respect to smallpox, although vaccinia virus (the virus used for smallpox vaccination) utilizes the CCR5 receptor during T-cell infection (Rahbar et al. 2006), it is unknown whether the smallpox variola virus also utilizes this

receptor or how the *CCR5*-Δ*32* allele would affect smallpox infection or mortality rates. Genetic drift (Sabeti et al. 2005) and balancing selection (Bamshad et al. 2002) have also been proposed to explain present day *CCR5*-Δ*32* allele frequencies.

Critical to each of these models is accurate dating of the *CCR5*-Δ*32* allele, but current methods for age estimation are burdened by large error ranges, typically on the order of thousands of years (e.g., Sabeti et al. 2005). Ancient DNA analysis is ideally situated to refine these estimates by providing direct genotype information at distinct periods in the past, which, when combined with historical records and radiocarbon dating, yield highly precise age information that allow a cemetery or skeletal assemblage to be dated to within a few hundred years. Moreover, when multiple time points are analyzed in concert, ancient DNA analysis can provide valuable information about allele frequency changes through time.

To date, three ancient DNA studies have reported *CCR5*-Δ*32* allele in European populations predating the 14th century CE: Dziekanowice, Poland, 1000-1100 CE (Zawicki and Witas 2008); Lichtenstein Cave, Germany, 900 BCE (Hummel et al. 2005); and Skateholm, Ire, Visby, Rossberge, Hjelmars ror, and Dragby, Sweden, 1690-5250 BCE (Lidén et al. 2006). Unfortunately, the Swedish samples exhibit evidence of allelic dropout, and thus calculation of genotype frequencies is problematic. The Polish and German samples appear well-preserved and exhibit divergent genotype patterns, with the Polish specimens indicating a lower *CCR5*-Δ*32* allele frequency one thousand years ago (4.4%) than today (10.3%) in Poland, while the German specimens exhibit a similar *CCR5*-Δ*32* allele frequency from the Bronze Age (11.8%) through the 14th century CE (approx. 12-14%) to the present (9.2%). Taken together, these results suggest regional differences in CCR5 genotype histories across Europe.

To independently test whether *CCR5*-Δ*32* allele frequencies have indeed remained unchanged in Germany since before the 14th century CE Black Death pandemic, we analyzed

tooth dentine from 53 individuals excavated from a 10^{th}-12^{th} century CE cemetery in central Germany. Using PCR-based techniques, we successfully generated CCR5 genotypes, amelogenin genotypes (genetic sex typing), and mitochondrial HVRI sequences for 32 individuals. Our results confirm that the *CCR5*-Δ*32* allele frequency in central Germany prior to the Black Death closely matches that found today in the same region.

Materials and Methods

The cemetery surrounding the St. Petri church and convent complex of Dalheim, Nordrhein-Westfalen, Germany dates to the medieval period (11^{th}-12^{th} century CE) and has been previously shown to be a good source of high quality ancient molecular data, both for historic human allele frequency studies (Kruttli et al. 2014) and for high-resolution analysis of ancient oral microbiomes (Warinner et al. 2014).

Within a dedicated ancient DNA clean-room and in accordance with established authenticity criteria as described in Warinner et al. (2014) and Kruttli et al. (2014), teeth from 53 individuals were surface cleaned with 1% hypochlorite solution, and the tooth roots were separated using a diamond bit rotary blade and ground to a fine powder using a freezer mill (SPEX SamplePrep). Approximately fifty milligrams of tooth powder were decalcified and digested overnight at 55˚C in a DNA extraction buffer consisting of 0.45 M EDTA and 10% proteinase K (Qiagen, >600 mAU/ml), followed by an additional 24-72 hours on a rotating nutator at room temperature until decalcification was complete. After centrifugation to pellet the cell debris, the supernatant was twice extracted with a 0.7 ml mixture of phenol, chloroform, and isoamyl alcohol (25:24:1) to remove proteins and lipids, followed by a final chloroform wash step. QiaQuick PCR purification columns were used to isolate DNA from the aqueous fraction, which was twice eluted into 30 ul of EB buffer following manufacturer instructions. At least two independent DNA extractions were performed for each individual, and one non-template control

extraction was performed in parallel for every six samples. The final concentration of 1 μl of extracted DNA was measured with a Qubit fluorometer high sensitivity assay (Qiagen). The non-template controls did not contain detectable amounts of DNA.

The resulting extracted DNA for each sample was first screened for: a) mitochondrial DNA (HVRI) to assess DNA survival and whether contamination had occurred, and b) nuclear DNA (amelogenin gene) to assess nuclear DNA survival and to genetically sex the individuals. Mitochondrial HVRI sequences and amelogenin alleles were PCR-amplified and analyzed by Sanger sequencing using an ABI 3730xl instrument (GATC Biotech) following previously described protocols (Kruttli et al. 2014). Next, a hydrolysis probe-based RT-PCR method was designed for detection of the *CCR5* wild type (wt) and 32-bp deletion (Δ*32*) alleles. This consisted of: Forward primer 5'- GGTCTTCATTACACCTGC-3'; Reverse primer 5'-TTTAGGATTCCCGAGTAG-3'; *CCR5*-Δ*32* probe 5'-(FAM)-TCTCATTTTCCATACATAAAGATAGTCATC-3'; and *CCR5*-*wt* probe 5'-(HEX)-TCAGTATCAATTCTGGAAGAATTTCCAG-3'. RT-PCR was performed on a Roche LightCycler 480 with the following cycling conditions: enzyme activation at 98˚C for 30 s, followed by 60 cycles of denaturation at 98˚C for 15 s and annealing and extension at 60˚C for 45 s. Wild type amplicons yielded a 140-bp product; *CCR5*-Δ*32* amplicons yielded a 108-bp product (primer inclusive). Following RT-PCR characterization, a subset of wt, Δ*32*, and wt/Δ*32* amplicons generated from nine individuals (B10, B60, B78, B88, G3, G6, G8, WM, and #a1) were cloned and sequenced to confirm the specificity of the assay.

All *CCR5* typing assays were performed a minimum of three times on two independent DNA extractions. Only samples with consistent results across at least three amplifications of both extracts were included for analysis. Any sample with evidence of multiple mtDNA profiles was excluded as a contamination precaution.

Results

Of the 53 individuals tested, 32 yielded high quality and consistent data for all three genetic loci (Table 1). Of these, 25 were homozygous *CCR5-wt* and seven were heterozygous (*CCR5-wt/CCR5-Δ32*). No samples were homozygous *CCR5-Δ32*. The resulting *CCR5-Δ32* frequency (0.109) confirms that the *CCR5-Δ32* allele was already common in central Germany before the 14th century CE Black Death pandemic, and it compares closely with modern frequencies reported for the nearby city of Gottingen, approximately 90 km to the west (0.098) (Hummel et al. 2005), and Mülheim, approximately 150 km to the east (0.106) (Lucotte and Mercier 1998). These results confirm that the *CCR5-Δ32* allele frequencies in this region of Germany are similar before and after the 14th century CE Black Death pandemic and do not support the hypothesis that strong selection has occurred at this locus during the past 700 years.

Discussion

The evolutionary history of the *CCR5-Δ32* allele is complex and likely regionally variable. Positive selection, balancing selection, and drift have each been proposed to explain current *CCR5* genotype frequencies across Eurasia, but there is currently no consensus on which scenario is most likely, nor when the allele most likely arose. Early hypotheses regarding recent positive selection related to the 14th century CE Black Death pandemic have been disproven by ancient DNA studies, but earlier viral epidemics, such as smallpox, remain candidates for possible earlier selection events. Unfortunately, too few ancient human genotypes are available to calculate accurate allele frequencies from relevant time periods for evaluating this hypothesis, and evidence of regional variability in the recent evolutionary history of *CCR5-Δ32* further complicates genetic modeling.

Previous ancient DNA studies suggested discordant rates of *CCR5-Δ32* allele frequency change in Central and Eastern Europe. Specifically, 14th century CE and present day allele

frequencies were reported to be similar in Germany, while *CCR5*-Δ*32* allele frequencies in Poland were reported to have doubled over the same period. To independently test whether *CCR5*-Δ*32* allele frequencies have indeed remained constant over the past millennium in Germany, we genotyped 32 individuals from a cemetery dating to the 10^{th}-12^{th} century CE and compared the allele frequencies to present-day populations within a 200 km radius. We found no difference in *CCR5*-Δ*32* frequencies between the two time periods, confirming the previous findings.

Importantly, in this study we used stringent criteria for genetic data inclusion, and we excluded poorly preserved samples that did not reliably amplify both mitochondrial and nuclear targets, a precaution not taken in previous ancient DNA studies. Moreover, we required multiple successful amplifications of the *CCR5* locus from independent PCR reactions and DNA extracts in order to make genotype calls. This is important because the fragmented nature of ancient DNA makes it highly susceptible to allelic dropout of longer targets, resulting in artificially high frequency estimates for the shorter allele. This amplification artifact is apparent in previously published Bronze Age data from Sweden, where only the *CCR5*-Δ*32* allele was detected in some assemblages (Lidén et al. 2006).

Further research is needed to better elucidate the evolutionary history of the *CCR5*-Δ*32* allele. Based on current studies, medieval *CCR5* allele frequencies appear to differ in Central and Eastern Europe, but it is unknown if this is related to selection, drift, or – in the case of the Polish samples – admixture with other populations. Sampling of additional time points is also necessary, particularly before and after the Bronze Age migration of Yamnaya populations (Haak et al. 2015). However, the difficulties of avoiding allelic dropout in amplification-based studies of insertion/deletion variants limit the application of conventional PCR techniques to older or poorly preserved skeletal assemblages. High-throughput sequencing in combination with targeted

capture-enrichment techniques present a promising alternative approach for recovering allelic frequencies from degraded samples. The analysis of ancient DNA from well-dated skeletal remains is a powerful method for testing competing hypotheses regarding the evolutionary history of the *CCR5*-Δ*32* allele.

Literature Cited

Bamshad, M. J., S. Mummidi, E. Gonzalez et al. 2002. A strong signature of balancing selection in the 5' cis-regulatory region of CCR5. *Proc. Natl. Acad. Sci. USA* 99(16):10539-10544.

Dean, M., M. Carrington, C. Winkler C et al. 1996. Genetic restriction of HIV-1 infection and progression to AIDS by a deletion allele of the CKR5 structural gene. *Science* 273(5283):1856-1862.

Elvin, S. J., E. D. Williamson, J. C. Scott et al. 2004. Evolutionary genetics: Ambiguous role of CCR5 in Y. pestis infection. *Nature* 430(6998):417.

Galvani, A. P. and M. Slatkin. 2003. Evaluating plague and smallpox as historical selective pressures for the CCR5-Delta 32 HIV-resistance allele. *Proc. Natl. Acad. Sci. USA* 100(25):15276-15279.

Hedrick, P. W. and B. C. Verrelli. 2006. "Ground truth" for selection on CCR5-Delta32. *Trends Genet.* 22(6):293-296.

Haak, W., I. Lazaridis, N. Patterson et al. 2015. Massive migration from the steppe was a source for Indo-European languages in Europe. *Nature* 522: 207-211.

Huang, Y., W. A. Paxton, S. M. Wolinsky et al. 1996. The role of a mutant CCR5 allele in HIV-1 transmission and disease progression. *Nat. Med.* 2(11):1240-1243.

Hummel, S., D. Schmidt, B. Kremeyer et al. 2005. Detection of the CCR5-Delta32 HIV resistance gene in Bronze Age skeletons. *Genes Immun.* 6(4):371-374.

Kruttli, A., A. Bouwman, G. Akgul et al. 2014. Ancient DNA analysis reveals high frequency of European lactase persistence allele (T-13910) in medieval central europe. *PLoS One* 9(1):e86251.

Lidén, K., A. Linderholm, and A. Götherström. 2006. Pushing it back. Dating the CCR5-D32 bp deletion to the Mesolithic in Sweden and its implications for the Meso/Neo transition. *Documenta Praehistorica* 33:29-37.

Liu, R., W. A. Paxton, S. Choe et al. 1996. Homozygous defect in HIV-1 coreceptor accounts for resistance of some multiply-exposed individuals to HIV-1 infection. *Cell* 86(3):367-377.

Lucotte, G., and G. Mercier. 1998. Distribution of the CCR5 gene 32-bp deletion in Europe. *J. Acquir. Immune Defic. Syndr. Hum. Retrovirol.* 19(2):174-177.

Martinson, J. J., N. H. Chapman, D. C. Rees et al. 1997. Global distribution of the CCR5 gene 32-basepair deletion. *Nat. Genet.* 16(1):100-103.

Mecsas, J., G. Franklin, W. A. Kuziel et al. 2004. Evolutionary genetics: CCR5 mutation and plague protection. *Nature* 427(6975):606.

Michael, N. L., G. Chang, L. G. Louie et al. 1997. The role of viral phenotype and CCR-5 gene defects in HIV-1 transmission and disease progression. *Nat. Med.* 3(3):338-340.

Rahbar, R., T. T. Murooka, A. A. Hinek et al. 2006. Vaccinia virus activation of CCR5 invokes tyrosine phosphorylation signaling events that support virus replication. *J. Virol.* 80(14):7245-7259.

Sabeti, P. C., E. Walsh, S. F. Schaffner et al. 2005. The case for selection at CCR5-Delta32. *PLoS Biol* 3(11):e378.

Samson, M., F. Libert, B. J. Doranz et al. 1996. Resistance to HIV-1 infection in caucasian individuals bearing mutant alleles of the CCR-5 chemokine receptor gene. *Nature* 382(6593):722-725.

Sharp, P. M. and B. H. Hahn. 2011. Origins of HIV and the AIDS pandemic. *Cold Spring Harb. Perspect. Med.* 1(1):a006841.

Stephens, J. C., D. E. Reich, D. B. Goldstein et al. 1998. Dating the origin of the CCR5-Delta32 AIDS-resistance allele by the coalescence of haplotypes. *Am. J. Hum. Genet.* 62(6):1507-1515.

Warinner, C., J. F. Rodrigues, R. Vyas et al. 2014. Pathogens and host immunity in the ancient human oral cavity. *Nat. Genet.* 46(4):336-344.

Zawicki, P. and H. W. Witas. 2008. HIV-1 protecting CCR5-Delta32 allele in medieval Poland. *Infect. Genet. Evol.* 8(2):146-151.

Table 1. HVRI variants, genetic sex, and CCR5 genotypes for ancient individuals and laboratory personnel

Sample	mtDNA HVRI SNPs (16000+)[a]	Sex Assay	*CCR5* *wt*/total	*CCR5* Δ*32*/total	*CCR5* genotype
Ancient samples					
296	176	XY	8/15	0/15	wt
314	256, 270	XY	9/9	0/9	wt
B07	rCRS	XY	7/15	0/15	wt
B10[b.c]	rCRS	XY	5/30	15/30	wt/Δ32
B11	rCRS	XY	6/6	0/6	wt
B14	189, 362	XY	7/9	0/9	wt
B15	rCRS	XY	8/18	0/18	wt
B26a	rCRS	XX	6/6	0/6	wt
B27	189	XX	6/6	0/6	wt
B30b1	rCRS	XX	8/12	0/12	wt
B30b2	rCRS	XX	6/12	0/12	wt
B34	192	XX	10/21	0/21	wt
B36	368	XY	6/15	6/15	wt/Δ32
B39	093, 362	XY	7/12	0/12	wt
B40	189	XY	8/9	0/9	wt
B41	222, 224, 311	XX	6/9	0/9	wt
B50	186, 189	XY	9/12	0/12	wt
B52	rCRS	XX	8/21	0/21	wt

B57	192	XX	6/15	6/15	wt
B59[d]	rCRS	XY	15/21	1/21	wt
B60[c]	192, 256, 270	XX	5/15	7/15	wt/Δ32
B78[c]	rCRS	XX	5/9	4/9	wt/Δ32
B85	163	XX	7/9	0/9	wt
B85a1	354	XX	6/6	0/6	wt
B88[c]	rCRS	XX	3/9	4/9	wt/Δ32
G03[c]	294, 296, 304	XY	9/21	0/21	wt
G06[c]	rCRS	XX	7/21	0/21	wt
G08[c]	189, 258, 265, 282	XY	3/27	7/27	wt/Δ32
G12	356, 362	XX	7/9	0/9	wt
#a1[c]	rCRS	XY	9/15	0/15	wt
#b	293, 311	XY	11/12	0/12	wt
WM[c]	298	XX	9/12	7/12	wt/Δ32
Laboratory personnel					
I1	172	XX	3/3	0/3	wt
I2	362	XX	3/3	0/3	wt
I3	051, 93, 184, 189, 234, 342	XX	3/3	0/3	wt
I4	182, 183, 189, 234, 342	XX	3/3	0/3	wt

Notes:

[a]SNPs are reported relative to the revised Cambridge reference sequence (rCRS). HVR1 sequences without SNPs are reported as rCRS.

[b]Three DNA extractions were analyzed.

[c]Amplicons from this individual were sequenced to confirm assay specificity. DNA sequences were consistent with the genotypes identified by RT-PCR.

[d]Because the deletion was observed only once and could not be replicated, it was excluded from analysis.